

Southern Sudan Interactive Radio Instruction (SSIRI)

Grade 1 Evaluation

May, 2008



TABLE OF CONTENTS

LIST OF ACRONYMS	II
EXECUTIVE SUMMARY	III
1.0 BACKGROUND	1
1.1 DESCRIPTION SSIRI AND THE SOUTHERN SUDAN CONTEXT	1
1.2 IRI ENROLMENT AND PARTICIPATION	1
2.0 EVALUATION METHODOLOGY.....	3
2.1 PURPOSE OF THE EVALUATION	3
2.2 PRETEST SAMPLE	3
2.3 POSTTEST SAMPLE	4
2.4 TEST DEVELOPMENT	4
<i>Test planning</i>	5
<i>Test Construction</i>	5
2.5 DATA COLLECTION, ENTRY AND ANALYSIS	6
<i>Training of Test Administrators</i>	6
3.0 FINDINGS	7
3.1 GRADE 1 ACHIEVEMENT IN ENGLISH, MATHEMATICS AND LOCAL LANGUAGE LITERACY	7
3.2 USE OF IRI AND CONTEXTUAL FACTORS THAT PROMOTE IRI EFFECTIVENESS	10
4.0 DISCUSSION AND RECOMMENDATIONS.....	13
4.1 ACHIEVEMENT BASIC SKILLS IN ENGLISH, LOCAL LANGUAGE LITERACY AND MATHEMATICS.....	13
4.2 RECOMMENDATIONS.....	15

LIST OF ACRONYMS

AMURT	Amanda Marga Universal Relief Team
EDC	Education Development Center, Inc.
IDPs	Internally Displaced Persons
IRI	Interactive Radio Instruction
M&E	Monitoring and Evaluation
MoEST	Southern Sudan Ministry of Education, Science and Technology
NGO	non-governmental organization
NRRDO	Nuba Relief, Rehabilitation and Development Organization
NSCSE	New Sudan Centre for Statistics and Evaluation
PST	Professional Studies for Teachers
REACH	Radio Education for Adults and Children
SBEP	Sudan Basic Education Programme
SSIRI	Southern Sudan Interactive Radio Instruction
TERBIA	Teaching English through Radio-Based Instruction for All
TERBIA A	Terbia Advanced Level
TERBIA B	Terbia Basic Level
TTI	Teacher Training Institute
USAID	United States Agency for International Development
USAID/SFO	USAID/Sudan Field Office

EXECUTIVE SUMMARY

Southern Sudan Radio Instruction Program (SSIRI) is a radio-based learning intervention to increase access to quality education opportunities and to improve teaching at the lower primary level in Southern Sudan. The P1 IRI programs were first broadcast in 2005. Programs address Southern Sudan curriculum goals for English, Local Language Literacy, and Mathematics. In addition to P1, programs have been introduced in P2 through P4 classes. This report presents findings of the first evaluation of learning achievement at the P1 level. The evaluation answers questions on whether learners that are exposed to IRI programs achieve basic skills in literacy and numeracy as expected for the Primary 1, the extent to which teachers use the IRI methodology as stipulated in the programs, as well as the contextual factors that enhance the effectiveness of IRI.

To answer evaluation questions, data was collected from a sample of 49 schools with 738 learners selected to represent 376 IRI participating schools. For learning assessment, a curriculum-based achievement tests was administered two times, a pretest in April and a posttest in November 2007. Findings summarized below respond to specific evaluation questions and provide insights into the characteristics of the teachers and learners in governments schools in Southern Sudan.

IRI participating schools in Southern Sudan enrolled 42,045 learners, 62.4 percent boys and 37.6 girls for all levels of primary in 2007. With 53.8 percent of P1 learners being older than 8 years, learners were generally older than expected for the P1 grade. Most teachers (88.2 percent) in IRI classes have a formal teaching qualification; most of them (60.0 percent) have also been trained in IRI.

The first question of the evaluation was whether learners that are exposed to IRI programs achieve basic literacy and numeracy skills stipulated in the curriculum for the Primary 1 level. Significant gains were registered in all three subtests, even though learners performed poorly in Local Language Literacy. Learners registered a gain of 27.8 percent in English, from a mean pretest of 15.4 percent to a mean posttest of 43.2 percent. Learners gained 12.7 percent in local language literacy, from the mean pretest of 15.3 to 27.7 percent in the posttest, while the gain in mathematics was 21.8 percent (from 35.1 percent to 56.8 percent).

The results showed, also, that learners in IRI schools performed better in all three subtests than learners in non-IRI control schools. The mean difference between IRI and control learners was 14.1 percent for English, 4.3 percent Local Language Literacy and 4.3 percent for Mathematics. There were benefits in learning for learners who attended school and were present during IRI lessons, and for learners whose teachers were trained in IRI performed better than those whose teachers did not receive IRI training.

While several suggestions for improvement were made throughout the report, recommendations and follow-up action for project improvement are as follows:

1. There is need to set performance standards or targets for each grade level. This will be useful in providing guidance on what the critical learning behaviors and outcomes for each grade level ought to be. Besides reporting progress based on curriculum standards (and not mean

scores), performance standards will be useful for benchmarking the Southern Sudan education system against global standards in the education sector, as well as for tracking performance in the different regions of the country.

2. MoEST needs to clarify what 'local language literacy' is in the context of Southern Sudan and why it is important to learn a local language if it is not synonymous with teaching children in their 'mother tongue.' SSIRI could then come up with improved strategies for teaching local language literacy, as well as for MoEST to develop a comprehensive area/local language policy and tools for its implementation.
3. Field monitors should systematically investigate why the programs are not being used. A good place to start would be to ensure that there is good reception in all areas, that broadcast times are published in all states, and that Outreach Coordinators assist teachers with timetabling issues where those arise.

1.0 BACKGROUND

1.1 Description SSIRI and the Southern Sudan context

Southern Sudan Interactive Radio Instruction (SSIRI) is a project that uses radio-based learning and other technologies to expand non-formal and alternative education opportunities for Southern Sudanese people. Its goal is to support efforts to improve access and quality in education in Southern Sudan and the three regions of Abyei, Blue Nile, and Southern Kordofan. With funding from USAID, SSIRI is jointly implemented by Southern Sudan Ministry of Education, Science and Technology (MoEST) and Education Development Center, Inc. (EDC). SSIRI is one of several USAID funded programs that were designed to help reconstruct an education sector which was crippled by many years of civil war. According to estimates by the New Sudan Centre for Statistics and Evaluation (NSCSE) and UNICEF, Southern Sudan had a gross enrolment ratio of 25.3 percent at the primary school level in 2004, which is the lowest access to primary education in the world¹. At 35.0 percent, Southern Sudan also had the lowest ratio of female to male enrolment.

Other key findings of the NSCSE investigation indicated low survival rates to P5 (28 percent) and even lower rates at P7 (2 percent). Only 6 percent of teachers received at least one year of pre-service training, 45 percent with two weeks to three months of in-service teacher training, and 49 percent of teachers have no training at all. Only 7 percent of the country's teaching force are women. The adult literacy rate is just 24 percent, and even lower for women at 12 percent. . Given the largely untrained teaching force, learners experience a wide range of pedagogies and teaching practices with such variations and inconsistencies presenting further obstacles to learning. In 2005 the Ministry of Education, Science, and Technology (MoEST) announced its aims to (a) enroll 200,000 additional pupils to reach a total of 700,000 pupils, and (b) to recruit 8,000 more teachers by the end of 2006. As ambitious as this goal may seem, it still left more than half of Southern Sudan's 1,500,000 school-aged children without access to quality education. Some of the issues mentioned here are at the heart of SSIRI and its IRI, TERBIA and PST elements.

The three components of SSIRI are *Learning Village*, a supplementary program to enhance teaching and learning of the primary school curriculum; Teaching English through Radio-Based Instruction for All (TERBIA)² that engages youth and adult learners in Civic Education and English Language instruction, and Professional Studies for Teachers (PST), a twelve-week accelerated in-service teacher training course. Project objectives that correspond to these components include improving access to learning opportunities, enhancing the primary program to attain higher learning achievement gains, improving teaching skills of participating teachers, and improving the capacity of government officials in management of radio programs.

1.2 IRI enrolment and participation

In the two years of implementation (2006 and 2007), SSIRI has achieved important milestones of developing P1 to P4 IRI programs, enlisting lower primary classes to participate in IRI, distributing radios and teachers guides to IRI participating schools, enlisting and contracting with radio stations

¹ Towards a baseline Best Estimates of Social Indicators in Southern Sudan, UNICEF and New Sudan Centre for Statistics and Evaluation in association with UNICEF, May 2004

² Teaching English through Radio-Based Instruction for All

to broadcast SSIRI programs, and training teachers in the IRI techniques. SSIRI lessons were being broadcast to 42,045 Primary 1 and 2 learners, 36,711 in 7 states of Southern Sudan, and 5,334 learners in the Three Areas (Blue Nile, Abyei, and Southern Kordofan) by the end of December, 2007. Table 1 presents learner enrolments by area/state and sex..

Table 1: Number of IRI learners in 2007, by state and sex

State	Female	Male	Total
Southern Sudan			
Western Equatoria	4 914	3 276	8 190
Jonglei	14 429	7 995	22 424
W. Bahr el Ghazal	462	308	770
Upper Nile	651	434	1 085
Central Equatoria	315	210	525
Lakes	1 222	640	1 862
Sub-Total	23 106	13 605	36 711
Three Areas			
Blue Nile	986	778	1 764
Southern Kordofan	378	252	630
Abyei	1 764	1 176	2 940
Sub-Total	3 128	2 206	5 334
Total	26 234	15 811	42 045

Approximately 2 of 3 learners in IRI participating schools (62.4 percent) are boys, while girls make only 37.6 percent, a slight improvement over the 35 percent participation rate of girls cited earlier. SSIRI has also made progress in training P1 and P2 teachers in the IRI methodology, a necessary condition for effective and efficient use the IRI programs. Table 2 indicates that a total of 828 teachers have been trained in IRI. Only 14.9 percent of the trained teachers are women.. Teachers who were trained in IRI have also been provided with radios and teachers guides.

Table 2: Number of teachers trained in IRI, by state and sex

Region	Male	Female	Total
Southern Sudan	643	109	752
Three areas	62	14	76
Total	705	123	828

SSIRI has a number of project activities that include developing programs, airing the lessons, distributing radios and other materials, training of teachers, and monitoring implementation of the programs. All these are routinely monitored by the M&E section. The focus of this evaluation was to assess the impact of IRI. The evaluation set out to investigate the extent to which IRI was implemented in Southern Sudan and the Three Areas, and whether or not IRI participating teachers were maximizing the opportunity provided by the programs by using IRI as prescribed in their training. A sample of P1 pupils who had listened to the IRI programs (IRI learners) and those who were not using radio programs (control learners) were selected and subjected to learning achievement test in Mathematics, English, and Local language Literacy to investigate if IRI was an effective tool for learning. This report presents results of the evaluation, an interpretation of the findings, and recommendations for further action by the MOEST and/or SSIRI.

2.0 EVALUATION METHODOLOGY

2.1 Purpose of the Evaluation

The purpose of this evaluation is to investigate whether P1 learners exposed to IRI programs achieved basic numeracy, basic English language comprehension skills, and functional literacy in a local language as stipulated in the curriculum. The evaluation further investigated whether teachers are using the IRI programs as expected in terms of their interaction with the radio lesson, their interaction with learners, and using IRI materials such as teachers' guides. Factors that enhance the effectiveness of IRI as attendance of IRI lessons, and the learning environment were also investigated. To that end, the evaluation focused on the following questions:

1. Do learners that are exposed to IRI programs achieve basic skills in literacy and numeracy as expected at the Primary 1?
2. To what extent do teachers use the IRI methodology as stipulated in the programs?
3. What are the contextual factors (learner, teacher, or learning environment) that enhance the effectiveness of IRI?

2.2 Pretest sample

The population of IRI learners for which data was received was 42,045 learners attending 328 schools. The design of the evaluation study was pretest posttest with control groups. A multi-stage purposive sampling strategy was used in both the pretest and posttest. Factors that were considered in the selection of evaluation sites included:

1. Whether IRI was implemented in the area, in terms of radio reception and actual tuning into radio broadcasts by teachers
2. Donor's emphasis on the Three Areas, as well as urban areas,
3. Practical and logistical considerations such as accessibility of sites, flight availability, availability of test administrators, financial resources, availability of accommodations and transport.

The Learning Village is currently being implemented in 17 locations, 5 locations where implementation is done through partnership, and 12 locations where SSIRI project staff oversee implementation. The program reaches 376 schools and at least 376 P1 classes. Some schools have more than one P1 stream. Four (4) of 10 states in Southern Sudan were selected. Two (2) of the Three Areas were also selected, as well as 1 of the 3 major towns that met the criteria above. This gave adequate representation of all the areas where the program is currently being implemented.

The towns/areas of Maridi, Panyagor, Malakal, Kauda, Aweil East were selected as pretest sites. In Aweil East IRI was implementation through AMURT, a local NGO. These locations were stratified along urban/rural, the Three Areas, and the larger Southern Sudan. For the urban areas, Malakal was selected; Kauda represented the Three Areas; and Maridi, Aweil and Panyagor represented the greater Southern Sudan. The table below shows the classes targeted for the assessment. . The following table shows the sample calculation for classrooms used for testing P1 literacy, English, and numeracy.

Table 3: Pretest sample from IRI participating schools, by state/location

Location	Total P1 classes	IRI classes	Control classes	Total classes	Total Learners
Maridi	31	6	1	7	101
Panyagor	63	12	3	15	174
Malakal	20	4	1	5	78
Aweil	53	10	2	12	181
Kauda	18	3	1	4	30
Abyei	12	2	1	3	45
Total	197	38	8	46	609

In each location there are IRI schools where implementation is carried by SSIRI staff, and other schools where implementation is supported by partner NGOs. Both types of schools were included in the sample. At the school level, a random sample of 15 learners was selected. As far as possible, the learners for the posttest who have attended at least 80 percent of the lessons.

2.3 Posttest sample

To the extent possible, the posttest was administered in the same schools that participated in the pretest, with the intention also to select the pupils who participated in the pretest. Two locations were substituted during the post test, Malakal and Kauda as shown in Table 4.

Table 4: Posttest sample of schools, by state

Location	IRI classes	Control classes	Total classes	Total Learners
Maridi	6	1	7	99
Panyagor	12	3	15	227
Malakal	4	1	5	73
Aweil	10	2	12	184
Abyei	2	1	3	50
Juba	5	2	7	105
Total	39	10	49	738

First, Malakal was replaced with Juba. Malakal schools were closed during the posttest due to flooding and an outbreak of meningitis. Juba was a good location in that IRI broadcasts through FM radio that makes the signal stronger and clearer than most areas. The second location that was substituted was Kauda in Nuba Mountains. Schools in this area started their broadcast much later than other schools due to the late start of their school year. An equal number of schools that were pretested in Kauda were redistributed to Juba and Maridi.

2.4 Test Development

This section describes the rationale for developing a Local Language Literacy, Mathematics and English Language achievement tests. The tests developed were curriculum-based mastery tests. Stages of the test development process included test planning, item writing, and pilot testing. The test administration procedure is also described briefly.

Test planning

The test development process commenced by a content analysis for Primary 1, 2 and 3, performed by the Education Advisor and Test Development Specialist. Instructional objectives from the Mathematics, Local Language and English syllabi were analyzed with the intent to distinguish between developmental and terminal objectives. Three test plans were developed. In the absence of grade-level reading lists that usually indicate the reading levels of learners, the teachers' guide was particularly useful in that it specifies new English language and local language vocabulary and the numeracy skills that are presented in each lesson. The syllabus and teachers guides also guided the test development process in terms of the cognitive skills that children have to master at this formative stage of being introduced to formal learning.

Test Construction

The purpose of the test was to assess and evaluate if learners have mastered basic literacy skills in a local language as well as basic numeracy skills, and whether they could understand simple communication in English at the end of the Primary 1 syllabus. The guiding principles during test development was that assessment procedures should match the intentions of each learning target, hence, the behaviors elicited from learners included recalling certain facts, as well as performing certain tasks. For instance, the intention of the learning targets on language during the early stages of learning is that learners should comprehend language and begin to produce simple language. Their comprehension of language in the lessons is demonstrated by the acting out in response to simple instructions, hence, the assessment of language skills comprised mainly of requesting them to perform actions when given simple instructions.

One form of the test was constructed for each of the three learning areas for Primary 1. Where possible, a set of parallel items was presented from which the test administrator would select the item to present to the learner. Table 5 presents a summary of the skills assessed and the weighting of each skill area.

Table 5: Skill areas and corresponding test items for Primary 1 Test, 2007

Skill Area	Intended Learning Target	Tasks	Points
English Language	Recalling names	1, 2	4
	Simple comprehension of language	7	2
	Production of language (speaking)	3,4,5,6	11
Local Language Literacy	Production of language (speaking)	5	2
	Production of language (reading)	1, 3, 6	9
	Production of language (writing)	2, 4, 7	8
Mathematics	Counting and writing numbers	1, 2, 5, 6	11
	Number operations	3, 4, 7, 8	8
	Naming and drawing shapes	9, 10	4

The English and Local Language tests each had 7 tasks, while mathematics was comprised of 10 tasks. All tests were developed in English, but only the English subtest was to be administered to the learner in English. Test administrators were required to translate the Local Language Literacy and Mathematics subtests from English to the area local language, presumed also to be the medium of instruction for schools in the area.

trial Testing

Items were pilot tested by test administrators in two centers in Rumbek. Trial testing assessed

whether the questions elicited the intended behavior/skills, and whether the correct difficulty levels in terms of content and language were maintained. The amount of time it took to administer the test was important in that children at the P1 age have a short attention span, and whether the proposed administration procedure was reasonable. After trial testing, a debriefing session was convened to receive additional feedback on how the test functioned. Interactions between learners, test administrators, and the test were noted for interpretation and for improving the test. Explanatory notes for each test form and a quick reference guide for the test administrators were developed. Trial testing provided feedback on the reasonableness and appropriateness of the test for testing literacy and numeracy skills at P1, and whether the learners were able to handle the format of the test.

2.5 Data Collection, entry and analysis

Training of Test Administrators

Training of test administrators was conducted in Rumbek by the SSIRI technical team, using a test administration booklet. Test administrators were briefed on the purpose of test, how the test was developed, the behaviors that each item intended to elicit, and how it was to be scored. Test administrators practiced administering the test in pairs, and then went out to the schools to conduct trial tests.

Live Testing

The pretest was administered close to the beginning of the school year in April 2007 over a period of 15 working days, while the posttest was conducted in November 2007 for another period of 15 working days. There were 5 teams of test administrators, each consisting of 3 people. Each of the team members was assigned a specific responsibility in the test administration.

Data Entry

Data entry commenced soon after the testing. Two data entry assistants entered the data into MS-access. Data was transported into MS Excel and finally into SPSS for analysis and developing result tables.

3.0 FINDINGS

The design of the evaluation study was pretest posttest with control groups. The original sample of learners that was selected from seven locations, namely, Maridi, Panyagor, Aweil Abyei, Kauda and Malakal. Pretests were administered at these locations at the beginning of the year. However, Kauda was later dropped from the sample because the school year starts much later than in other areas and, thus, they would have covered fewer than half the broadcasts at the time of posttesting. Malakal schools were closed during the posttest due to flooding and an outbreak of meningitis. Juba was chosen as a replacement for the two locations. Juba represents an urban town and has an FM station with IRI broadcasts that has a clear signal. A total of 57 schools participated in both pretest and posttest. Forty-six (46) schools participated in the pretest. Sixteen (16) of the schools were dropped at the posttest and replaced by 11 schools. Hence the posttest sample comprised of 41 schools, 31 being IRI and 10 being control schools.

3.1 Grade 1 achievement in English, Mathematics and Local Language literacy

Performance on the pretest and posttest

The first question of the evaluation was whether learners that are exposed to IRI programs achieve basic skills in literacy and numeracy as expected at the Primary 1 level. The mean pretest for IRI learners was 15.4 percent in English, while the posttest score was 43.2 percent. Learners registered a gain of 27.8 percent. Learners gained 12.7 percent in local language literacy, from the mean pretest of 15.3 to 27.7 percent in the posttest. In mathematics the pretest score was 35.1 percent, while the posttest scores was 56.8 percent as shown in Table 6.

Table 6: Pretest and posttest means for all IRI learners, by subtest

Subtest	N	Maximum Score	Mean	Mean Percent	Mean Gain (%)
English Pretest	415	17.0	2.6	15.4	27.8
English Posttest	415	17.0	7.4	43.2	
Local Lang Pretest	418	19.0	2.9	15.2	12.7
Local Lang Posttest	418	19.0	5.3	27.9	
Mathematics Pretest	419	23.0	8.1	35.1	21.8
Mathematics Posttest	419	23.0	13.1	56.8	

The pretest scores were higher in mathematics than in than in English or Local Language Literacy. This is typical of performance at the Grade 1 level, mainly because children acquire more numeracy skills from non structured and non deliberate learning than they do with literacy skills. A paired-samples comparison indicates that differences in the pretest and posttest are significant,³ an indication that after a year of schooling children did benefit from the learning activities including IRI. IRI methodology deliberately sets out to make learning interactive and interesting, and to motivate all children to attend and participate.

³ English Pretest/Posttest: $t=20.41$, $p=.00$; Local Language Literacy Pretest/Posttest: $t=11.60$, $p=.00$; Mathematics Pretest/Posttest: $t=14.85$, $p=.00$

Comparison of IRI learners and a non-IRI control group

Some schools in the control sample were replaced, which means that the pretest and posttest control samples are not equivalent. However, this analysis treated the control samples as a random group of learners and compared the posttest scores with the pretest scores. Control learners performed significantly better than IRI learners in all three pretests as shown in Table 7.

Table 7: Pretest means, by type of learner and subtest

Learners	Subtest	N	Maximum Possible	Mean	Mean Percent	Mean Diff (%)
English Pretest	IRI Learners	535	17.0	2.6	15.4	-11.4
	Control Learners	60	17.0	4.6	26.8	
Local Lang Pretest	IRI Learners	546	19.0	2.9	15.2	-15.2
	Control Learners	60	19.0	5.8	30.4	
Mathematics Pretest	IRI Learners	539	23.0	8.1	35.1	-15.0
	Control Learners	60	23.0	11.5	50.1	

The deficit was 11.4 percent for English, and as high as 15 percent for Local languages and Mathematics. A sample of learners, most of which took the pretest, participated in the posttest at the end of the school year. The posttest sample comprised of 41 schools, 31 schools that participated in IRI, and 10 control schools. Table 8 below compares performance of IRI learners on the posttest with that of control learners.

Table 8: Posttest means, by type of learner and subtest

Learners	Subtest	N	Maximum Possible	Mean	Mean Percent	Mean Diff (%)
English Posttest	IRI Learners	585	17.0	7.4	43.2	14.1
	Control Learners	103	17.0	4.9	29.1	
Local Lang Posttest	IRI Learners	585	19.0	5.3	27.9	4.3
	Control Learners	104	19.0	4.5	23.6	
Mathematics Posttest	IRI Learners	586	23.0	13.1	56.8	4.3
	Control Learners	103	23.0	12.1	52.5	

In English IRI learners had a posttest mean of 43.2, while the learners in the control group posted a mean of 29.1 percent. The mean posttest for IRI learners was 27.9 percent in Local Language Literacy, compared to a posttest score of 23.6 percent for control learners. In mathematics the IRI posttest score was 56.8 percent compared to a posttest score of 52.5 percent for control learners. The difference between IRI and control learners was wider in English, where IRI learners seem to have benefited the most. An independent samples t-test comparison of means between IRI and control learners yielded significant differences⁴, an indication of the ‘value-added’ by IRI.

There were no significant differences between boys and girls in the pretest. Boys and girls posted similar performance on the posttest as shown in Table 9.

⁴ English Posttest: $t=6.37$, $p=.00$; Local Language Literacy Posttest: $t=3.56$, $p=.00$; Mathematics Posttest: $t=2.49$, $p=.01$

Table 9: Posttest mean scores by type of learners and sex

Type of Learner	Sex of Learner	Subtest	N	Maximum possible	Mean	Mean Percent
IRI Learners	Male	English Posttest	329	17.0	7.4	43.6
		Local Literacy Posttest	329	19.0	5.5	28.8
		Mathematics Posttest	330	23.0	13.5	58.9
	Female	English Posttest	256	17.0	7.3	42.8
		Local Literacy Posttest	256	19.0	5.1	26.7
		Mathematics Posttest	256	23.0	12.5	54.2
Control Learners	Male	English Posttest	46	17.0	4.9	28.8
		Local Literacy Posttest	47	19.0	4.2	21.9
		Mathematics Posttest	56	23.0	12.4	53.7
	Female	English Posttest	57	17.0	5.0	29.3
		Local Literacy Posttest	57	19.0	4.8	25.0
		Mathematics Posttest	56	23.0	12.4	53.7

However, there was a significant difference between boys and girls participating in IRI in the mathematics posttest scores. Contrary to expectation and previous IRI results elsewhere, IRI implementation put boys at an advantage over girls.

Mean comparison by age of learner

A number of learners that participated were quite young, younger than the school going age of 7 years old. For IRI learners results generally showed an increase in performance as age increased in two of the subtests with the oldest age group performing best. However, the 7-9 age group performed better than the older children in Mathematics as indicated in Table 10 below, even though the difference is not significant.

Table 10: Posttest mean scores for IRI learners, by age category

Age categories	Subtest	N	Maximum possible	Mean	Mean percent
6 years	English	59	17.0	7.1	41.8
	Local Language Literacy	59	19.0	4.2	22.1
	Mathematics	59	23.0	11.9	51.7
7 -9 years	English	447	17.0	7.1	41.8
	Local Language Literacy	448	19.0	5.1	26.8
	Mathematics	447	23.0	13.3	57.8
10 and above	English	182	17.0	8.0	47.1
	Local Language Literacy	182	19.0	6.0	31.6
	Mathematics	183	23.0	12.8	55.7

The mean differences for English and Local Language Literacy were significant⁵. A similar analysis for the control group returned non-significant means for both English and Mathematics. The data seems to suggest that age has to be taken into consideration for future enrolment policy and practice in schools.

⁵ English: $F = 3.2$, $p = .04$; Local Language Literacy: $F = 8.5$, $p = .00$, Mathematics: $F = 2.3$, $p = .10$

Mean comparison by number of IRI lessons attended

Of the 713 IRI learners who participated in the posttest, 125 did not have data on the attendance variable. Thirty nine (39) learners (6.6 percent) posted high attendance, which means that they attended 80 of 100 lessons or more, while 139 learners (23.6 percent) posted medium attendance (60-79 lessons). 410 learners (69.7 percent) attended 59 lessons or fewer.

Table 11: Posttest mean scores for IRI learners, by lessons attended

Attendance	Subtest	N	Maximum possible	Mean	Mean percent
High 80-100 lessons	English	39	17.00	8.1	47.5
	Local language Literacy	39	19.00	5.3	28.1
	Mathematics	39	23.00	14.3	62.0
Medium 60-79 lessons	English	137	17.00	8.6	50.6
	Local language Literacy	137	19.00	5.7	30.1
	Mathematics	137	23.00	13.8	59.9
Low 0-59 lessons	English	440	17.00	6.9	40.4
	Local language Literacy	440	19.00	5.2	27.2
	Mathematics	440	23.00	12.7	55.3

The group that posted the lowest participation in IRI performed significantly lower in all 3 subtests as indicated in Table 11. With medium attendees performing better in English and Local Language Literacy, and high attendees performing better in Mathematics, the results suggested that a minimum of 60 lessons is what learners needed to do well in all three subtests.

3.2 Use of IRI and contextual factors that promote IRI effectiveness

Means comparison by IRI training

The majority of the teachers in the sample (80.4 percent) had completed primary or secondary schools schooling, but did not have a formal teaching qualification, hence, initial IRI training prepared them to better receive and utilize the radio programs. 73.7 percent of P1 teachers in IRI participating schools received face-to-face training on IRI techniques. Performance for all learners was disaggregated by whether or not their teachers had received IRI training. Table 12 indicates that learners in classes where teachers were trained in IRI performed significantly better in English, and were at par with control learners in a third subtest. The mean differences were significant in English and mathematics.

Table 12: Posttest means for all learners, by training of teachers in IRI

IRI training	Subtest	N	Maximum	Mean	Mean Percent
Trained	English	541	17.00	7.3	42.8
	Local Language Literacy	541	19.00	5.1	27.1
	Mathematics	542	23.00	13.2	57.5
Not trained	English	147	17.00	5.9	35.0
	Local Literacy	148	19.00	5.3	27.9
	Mathematics	147	23.00	11.9	51.6

Broadcast Lessons missed

About two-third of the learners (63.5 percent) were in classes that missed eleven or more days of broadcasts due to radio problems according to teacher self-reports, while 36.5 percent were in classes that missed only up to ten days of broadcasts. Missing more than 11 days of broadcasts affected learner performance significantly for English and Mathematics as shown in Table 13.

Table 13: Posttest mean scores by broadcasts missed due to radio related problems

Lesson missed during the year		N	Maximum	Mean	Mean Percent
0-10 lessons	English	209	17.0	8.5	49.9
	Local Language Literacy	209	19.0	5.7	29.8
	Mathematics	210	23.0	14.1	61.2
11-20 lessons	English	166	17.0	6.8	40.1
	Local Language Literacy	166	19.0	5.9	31.0
	Mathematics	166	23.0	12.9	56.2
21 and over lessons	English	196	17.0	6.6	38.9
	Local Language Literacy	196	19.0	4.5	23.7
	Mathematics	196	23.0	12.2	53.2

Availability of Teacher's Guides

Performance was disaggregated by whether or not teachers had a teacher's guide for only part of the year or for the full year as indicated in Table 14. The results seem to suggest that the availability of the teacher's guide did not benefit the learners as was expected, but this may be due to the fact that teachers were only asked if they had the guide, and not if they used it.

Table 14: Posttest mean scores by availability of teacher's guide

Availability of teacher's guide		N	Maximum	Mean	Mean Percent
No guide	English	178	17.0	7.7	45.3
	Local Language Literacy	178	19.0	4.8	25.2
	Mathematics	178	23.0	13.1	56.8
Had guide since Term 1	English	177	17.0	6.1	36.1
	Local Language Literacy	177	19.0	4.7	24.6
	Mathematics	178	23.0	11.4	49.7
Had guide since Term 2	English	214	17.0	8.0	46.9
	Local Language Literacy	214	19.0	6.4	33.4
	Mathematics	214	23.0	14.3	62.2

Medium of Instruction

Teachers were asked about the medium of instruction used in their classrooms. Dinka was used as the medium of instruction for 49.2 percent of the learners, Juba Arabic was used for 28.0 percent of

the learners, while 21.1 percent of learners were reported to be in classes that used English as the medium of instruction. Table 15 indicates a definite benefit in English Language performance for children who used English as the medium of instruction. Other than that, learners who were taught in Dinka posted average performance in English and very good performance in Mathematics. Learners who were taught in Juba Arabic posted the weakest performance in each of the three subtests.

Table 15: Posttest mean scores by medium of instruction

Medium of instruction	Subtest	N	Maximum	Mean	Mean Percent
Dinka	English	330	17	8.1	47.4
	Local Language Literacy	330	19	5.9	31.0
	Mathematics	330	23	13.9	60.3
Juba Arabic	English	195	17	5.8	34.1
	Local Language Literacy	195	19	4.3	22.6
	Mathematics	195	23	11.9	51.9
English	English	60	17	8.5	50.2
	Local Language Literacy	60	19	5.4	28.2
	Mathematics	61	23	12.4	53.7

Location (state)

The data was also disaggregated by locality, and reported for the five states that were involved in assessment of learning. Table 15 reflects the posttest means for IRI learners by state. With a composite mean score of 53.5 for all three subtests Abyei state posted the highest performance while Western Equatoria performed lowest (a composite mean score of 32.3). Learners posted the strongest performance in Mathematics and weakest in Local Language Literacy in all 5 states.

Table 16: Posttest mean scores for IRI learners, by state

State	Subtest	N	Maximum Possible	Mean	Mean percent	Composite Mean
Abyei	English	46	17.0	9.3	54.9	53.5
	Local language Literacy	46	19.0	8.1	42.8	
	Mathematics	46	23.0	14.4	62.7	
Nothorn Bahr el Ghazal	English	150	17.0	7.7	45.1	46.2
	Local language Literacy	150	19.0	6.6	34.9	
	Mathematics	150	23.0	13.4	58.5	
Western Equatoria	English	133	17.0	5.3	30.9	32.3
	Local language Literacy	133	19.0	3.8	20.2	
	Mathematics	134	23.0	10.5	45.8	
Jonglei	English	163	17.0	8.2	48.5	43.8
	Local language Literacy	163	19.0	4.3	22.8	
	Mathematics	163	23.0	13.8	60.2	
Central Equatoria	English	93	17.0	7.3	42.9	44.5
	Local language Literacy	93	19.0	5.5	29.2	
	Mathematics	93	23.0	14.1	61.4	

While the overall percentage for low category learners was 69.7 percent, the proportion of low category learners in Western Equatoria was 79.9 percent, which means that they missed considerably

more IRI lessons. Participation in IRI lessons was lowest in Central Equatoria (85.9 percent in the low category), even though learners in the state performed better in Mathematics.

4.0 DISCUSSION AND RECOMMENDATIONS

The focus of the evaluation was to assess the impact of IRI on learning at Primary 1 in Southern Sudan. The evaluation answers questions on the extent to which teachers use the IRI methodology as stipulated in the programs, whether learners that are exposed to IRI programs achieve basic skills in literacy and numeracy as expected at the Primary 1, and whether there are contextual factors (learner, teacher, or learning environment) that enhance the effectiveness of IRI. Performance of IRI and control learners was also compared. This section discusses findings and makes recommendations that have implications for strengthening implementation of SSIRI.

4.1 Achievement basic skills in English, Local language Literacy and Mathematics

The results of the evaluation indicate that IRI had a significant impact on learning. IRI learners posted significant gains between the pretest and posttest in all three subtests. But if the mean score is considered as a percent of the curriculum content that was mastered, performance in Local Language Literacy was poor (27.9 percent), somewhat better in Mathematics (43.2 percent), and even better in English (56.8 percent). IRI learners registered the highest gains between the pretest and posttest in English. This was due mostly to the fact that IRI programs are produced and broadcast in English, even though there are Local language Literacy and Mathematics segments in each program. The Local Language Literacy and Mathematics segments are directly communicated to the class in English, and teachers are asked to translate into the local language. Hence, there is more exposure to the use of English language.

Conversely, there was no exposure or modeling of local language use from the radio teacher, perhaps accounting for the poor performance on the local language literacy test. Reasons postulated for poor performance in the local language test include the fact that the Local Language Literacy test was also developed in English, with test administrators being directed to read each item and translate it into the area local language, presumably the medium of instruction for the learners. Reports from the field indicate that the strategy did not work well; there were inadequacies in communicating the tasks to the learners in cases where test administrators were not sufficiently proficient in the area local language.

Second, the curriculum stipulates the teaching of ‘mother tongue’. Mother tongue in has been operationalized IRI programs to mean an ‘area local language’ or language of the catchment area.. But whatever the case, most schools do not teach in mother tongue or in the local languages beyond what IRI programs prescribe. There are no local language learning materials in most schools. This means that children do not get to see the written text, let alone acquire the necessary reading or writing skills in a local language. Where schools do make an effort to teach literacy in the local language, learners in many classes speak different languages, which means that what is presumed to be first language for all students may, in fact, be a second language for many of the students. Neither IRI nor the conventional teaching programs prepare teachers adequately for local language teaching, or to handle such complexities. A further investigation (see Table 15) revealed that 20 percent of the

learners use English as a medium of instruction, which suggest that there is no exposure to a Local Language for this group of learners.

While SSIRI's role is to make sure that IRI is implemented to the fullest for learners to derive maximum benefit, SSIRI can also assist in influencing MoEST towards putting in place or clarifying certain policy frameworks. First, there is need to set performance standards on learning targets. This will be useful in providing guidance on what the critical learning behaviors and outcomes for each grade level ought to be, as well as reporting progress based on curriculum standards (and not mean scores). Secondly, there is need to clarify what 'local language literacy' is in the context of Southern Sudan and why it is important to learn a local language if it is not synonymous with teaching children in their 'mother tongue.' This would help SSIRI to come up with strategies for teaching local language. To the extent that children in the earliest grades learn better in a language they are familiar with, the relevant authorities in MoEST should be assisted to develop a comprehensive area/local language policy and tools for its implementation.

Comparison of IRI and control groups

In comparing IRI and control learners, control learners performed significantly better in all pretests, a result that is difficult to explain since the project aimed at a random sample in both groups. The results indicate that the control sample had an advantage over the IRI sample at the beginning of the school year. Conversely, IRI learners performed significantly better than non-IRI control learners in all three posttests. IRI methodology deliberately sets out to make learning interactive and interesting. Using song, play and a variety of actors, learners are required to listen attentively and respond actively to different activities several times during the broadcast. Implicit in each lesson are teacher training strategies designed to motivate the learners and increase their chances of success for both teachers and learners.

For example, the radio actively engages teachers in organizing in a certain manner; by calling children to the front, asking children to work in pairs or small groups, work as a whole class, and to explore the learning environment. In addition, the lessons make a point of reviewing and reinforcing skills and concepts taught in earlier lessons and suggesting activities for before and after the broadcast. Teacher practices are also reinforced, and there is a likelihood of a transfer of these instructional techniques in other classroom settings.

Comparison by sex and age for IRI and control groups

Performance between boys and girls was similar in all pretest scores, and in all but one posttest scores. Boys performed significantly better than girls in the mathematics posttest. The IRI methodology has been shown to be successful in providing equal opportunity to boys and girls, and being an 'equalizer' in learning achievement in other places. SSIRI should examine their programming, teacher training, and actual classroom practices to ensure that girls are not being unduly disadvantaged in mathematics. Additional strategies should be found to make mathematics learning more accessible to girls so that they can succeed in a school culture where girls and women are a minority. One of the possible actions would be to increase role models for girls at the lower primary grades by bringing in more women teachers into the schools.

According to our sample, 40 percent of P 1 learners were older. This is common in most developing countries generally, and typical for post-conflict situations in particular. Children have to reach a

certain critical age to better cope with learning of concepts. Even though the correlation between age and performance was not perfect, the results indicated that from 7 years old children performed better; they were developmentally ready to engage in learning activities. SSIRI can avail this information to MoEST to assist in determining school readiness and the official school-going age.

Contextual factors that enhance IRI effectiveness

A number of factors are necessary for effective learning using IRI. A well functioning radio and clear signal are necessary. Also, teachers have to be trained in the IRI approach and they have to use the radio consistently. Both teachers and children have to participate as directed by the radio teacher and teachers and have to follow through with activities prescribed for the period before and after the broadcast. These and related issues were the subject of the questionnaire administered to teachers in IRI schools sampled for assessment. Learner performance was then disaggregated by factors such as training of teachers in IRI and whether teachers were using a teacher's guide.

Mean differences for learners whose teachers were trained in IRI were significantly better in English and Mathematics. There was no difference in Local Language Literacy in mean scores for learners whose teachers were trained in IRI (27.9 percent) and for learners whose teachers were not trained in IRI (27.1 percent). This finding further confirms the observations above on the need to revisit, not only the testing, but the policy on the teaching of Local Language Literacy. A more pressing issue for SSIRI is that all segments of IRI programming are supposed to demonstrate value added to of targeted learning areas, thus, adequacy of Local Language Literacy segments of IRI programs should be investigated systematically.

Locality was another contextual factor that yielded differences in performance. Of the five states that were sampled, learners from the Abyei state performed best in all three subtests (with a composite score of 53.5), while those from Western Equatoria posted the weakest with a composite score in the three subtests 32.3. Contrary to the findings, Western Equatoria was expected to perform better since the region is at a relative advantage in terms of being situated in a more reachable location, has good radio reception, and schools in the state continued to run even when there were interruptions in other places. As the state that hosts the SSIRI office within Southern Sudan, Central Equatoria also performed below expectations.

From self reports of teachers, an aggregate of 69.7 percent in all five states fell in the low attendance category (which means that they were exposed to 59 lessons or fewer). We found from self-reports of teachers that at 80 percent, Western Equatoria had the second highest proportion of learners in the low IRI attendance category, a factor that may largely explains low performance in the state. Low participation rates in any instructional initiative undermine its effectiveness. At the very least, there is a serious implementation threat for SSIRI if teachers report that they are not using the programs. Field monitors should systematically investigate why the programs are not being used. A good place to start would be to ensure that there is good reception in all areas, that broadcast time schedule is published in all states, and that Outreach Coordinators assist teachers with timetabling issues where those arise. In addition, the SSIRI M&E section should insure that an agreed number of field monitoring reports are received from monitors and verified every month

4.2 Recommendations

The discussion above posted a number of suggestions for improvement; some of these are recommended for follow-up action by SSIRI, while those with implications for policy-making can be investigated further and followed up with MoEST.

1. There is need to set performance standards or targets for each grade level. This will be useful in providing guidance on what the critical learning behaviors and outcomes for each grade level ought to be, as well as reporting progress based on curriculum standards (and not mean scores). Performance standards are also useful for benchmarking the Southern Sudan education system against global standards in the education sector, as well as tracking performance in the different regions of the country.
2. MoEST needs to clarify what 'local language literacy' is in the context of Southern Sudan, why it is important to learn a local language if it is not synonymous with teaching children in their 'mother tongue.' SSIRI could then strengthen its strategies for teaching local language literacy and the MoEST could further develop a comprehensive area/local language policy and tools for its implementation
3. Field monitors should systematically investigate why the programs are not being used. A good place to start would be to ensure that there is good reception in all areas, that broadcast times are published in all states, and that Outreach Coordinators assist teachers with timetabling issues where those arise.